

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

The Future of Testing

Buros-Nebraska Series on Measurement and  
Testing

---

1986

## 9. New Directions for Interest Testing

John L. Holland

*Johns Hopkins University*

Follow this and additional works at: <https://digitalcommons.unl.edu/burosfuturetesting>

---

Holland, John L., "9. New Directions for Interest Testing" (1986). *The Future of Testing*. 10.  
<https://digitalcommons.unl.edu/burosfuturetesting/10>

This Article is brought to you for free and open access by the Buros-Nebraska Series on Measurement and Testing at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in The Future of Testing by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# 9

## New Directions for Interest Testing

John L. Holland

*Johns Hopkins University*

This chapter is an attempt to outline where interest testing may be or should be in the near future: What changes will be seen in the development or revision of inventories, what new areas of application will occur, and what technical, social, and professional problems need resolution to get to a more desirable future.

This sounds like a rational task. I have been asked to describe a desirable future by canoeing through the rapids of psychometric fashions, disgruntled test takers, passive publishers, worried professionals and their righteous associations, and future islands of unpredictable theory. To make this task easier, the sponsor cautioned me to rely on empirical data, not daydreams.

Fortunately, I can recognize an impossible task without the aid of consultants. For several reasons, it appears helpful to redefine the task. Earlier opinions by distinguished pioneers in interest measurement have occasionally been off the mark. For example, Kuder (1954) suggested that occupational titles made poor items and that activity items would be the wave of the future. News item: Occupational items continue to be useful and popular in most inventories. And inventories that use only activity items usually include occupational titles disguised as “Be an accountant,” or “Be a counselor.” Developers apparently get tired of looking for good items by following a restrictive rule. At an earlier time, Strong (1943) and others dismissed a person’s vocational aspiration as a weak index of the occupation a person would actually enter because this index did not have a substantial correlation with a person’s measured interests. However, Dolliver (1969) started a cottage industry of research by demonstrating that aspirations and measured interests have about equal predictive validity. Later, we learned that the use of interest inventories and aspirations in tandem produced very substantial predictions.

These events and the work of futurists imply that it is helpful to see future developments not only as the continuation of current trends but also as developments that will be shaped by economic, social, technological, and theoretical forces that we cannot always anticipate or control. Consequently, I attempt to relate current developments to future developments, but my forecasts will surely be deflected by unanticipated events. I also try to distinguish long-term trends that I believe are desirable and helpful and those that may be undesirable and not helpful.

My reservations about this task have considerable empirical support. I have multiple conflicts of interest. I am the author of two interest inventories that have been the object of close scrutiny for 10 years. I am familiar with the evidence and issues about inventory biases, development, effects and usage, but my beliefs about these matters have received only mixed reviews (Gottfredson & Holland, 1978; Tittle & Zytowski, 1978). The most constructive outcome of this experience for me has been to perceive interest inventories in the context of usefulness, validity and reliability—and about in that order.

### USEFULNESS, VALIDITY, AND RELIABILITY

Before I get to the main topic, it appears helpful to define usefulness and to indicate its relation to validity and reliability, for usefulness plays a key role in the evaluation of inventories and in their future development.

More explicitly, I view inventories as both interventions and assessment devices. They are interventions, because they are used to create one or more beneficial effects. Interest inventories can support a person's vocational aspiration, stimulate a comprehensive exploration of occupational possibilities, or provide a structure for understanding vocational interests and the occupational world.

They are also assessment devices, because they provide the test taker and counselor with a systematic account of a person's interests. Consequently, the validity and reliability of that information contributes to an inventory's usefulness.

Despite some interdependence of these constructs, it seems desirable to restrict validity and reliability to the data about an inventory's psychometric characteristics and to the literal interpretation of these data rather than to assume that these constructs can also account for all of an inventory's effects on the test-taker as well as its usefulness from the test-taker's point-of-view. Usefulness would then apply only to the effects of an inventory on the test-taker and to those aspects of an inventory that are assumed to enhance effects—clarity of language, influential interpretive materials, linkage to other interventions, and other sources of occupational information.

Some obscure research and development illustrate how loosely coupled these constructs—validity, reliability, and usefulness—can be. For instance, dissimilar interest inventories, which have been assumed to have divergent predictive validities, appear to have very similar beneficial effects on the test-taker (Pallas, Dahmann, Gucer, & Holland, 1983). Still other evaluations have indicated that many special experiences with obvious unreliable and invalid qualities such as interviewing an occupational representative or filling out a simple unresearched form about a particular vocational aspiration are frequently rated by test-takers as more helpful than taking the Kuder Occupational Interest Survey, the Strong-Campbell Interest Inventory, or the Self Directed Search (Evans & Rector, 1978; Rayman, Bernard, Holland, & Barnett, 1983).

These evaluations imply a need to perceive inventories in two contexts:

1. Does the information provided for decision making have at least a moderate degree of construct validity and retest reliability?
2. Are the effects of taking the inventory consistent with one or more test-taker goals?

This two-fold perspective appears helpful for understanding where we have been in interest measurement and where we appear to be going. It also provides a structure for resolving old and potential controversies about the biases or effects of inventories.

This perspective clarifies some issues by suggesting a more analytical and pragmatic approach to evaluation and inventory development. In the same way that we have learned that validity and reliability are ambiguous words that require elaboration into several elements, it is also helpful to divide usefulness into elements. In this instance, the most common uses are: (a) exploration or increasing the range of occupations a person will consider, (b) reassurance or providing support for a person's vocational aspiration or potential job, a goal that depends on an inventory's predictive validity, and (c) self-understanding or providing structured information for comprehending the character of one's interests. The uses or purposes of interest testing overlap those that Sundberg (1977) has proposed for personality assessment. He proposes "image forming," analogous to self-understanding; and "decision-making," analogous to reassurance and exploration. Sundberg also proposes "theory building" or the use of assessment data to develop theoretical concepts. Most people would consider "theory building" as a secondary goal of interest testing, although the data generated from interest inventories have stimulated some theoretical work.

With this perspective and my reservations, I outline the current status of interest assessment, some critical issues and developments, and finally, what the future may be or should be like.

## CURRENT STATUS OF INTEREST TESTING

The current status of interest measurement is characterized by great diversity. More than 200 interest inventories (Buros, 1978, pp. 1549–1649) have been developed by following a wide range of strategies and methods: homogeneous scaling, defined criterion groups, factor analysis, and other psychometric methods or theoretical models of vocational interests.

The diversity in inventory construction also is matched by an enormous increase in the diversity of test takers. Ten years ago, inventories were most popular among high school and college students along with scattered industrial applications—largely in research or selection. Now inventories are being used by adults of all ages in career, educational, retirement, and recreational planning. Some inventories have been adapted for use with special groups—notably the visually-impaired; many are in one or more foreign languages, and several interest inventories have been developed for particular groups. The most striking example is the development of inventories for people with skilled trades interests.

The current status of interest assessment is also characterized by a shift from the traditional preoccupation with an inventory's psychometric characteristics to a preoccupation with its therapeutic or not so therapeutic influences (Zytowski & Borgen, 1983). This shift has been stimulated by multiple social events, evaluative research, incidental psychometric findings, and hard times.

Until 1972, there were no experimental studies on the effects of taking an interest inventory. Developers and counselors were concerned primarily with construct validity but especially an inventory's predictive validity. At that time, inventories were imbedded in the information for decision-making model. The counselor's role was to help a person obtain an accurate interpretation of the information in the interest profile and to integrate that information with other personal and occupational data. Counselors and psychologists were typically more concerned with therapeutic techniques than with the independent influence of interest inventories.

The women's movement, the attacks on the Strong Vocational Interest Blank (SVIB) and the Self Directed Search (SDS), the Association for Measurement and Evaluation in Guidance Commission report (AMEG, 1973), and a National Institute of Education Conference (Diamond, 1975) stimulated a strong interest in the influence of interest inventories on women. This controversy eventually generalized to the effects of inventories on blacks and whites, poor and rich, and other groups and stimulated about 100 pieces of research performed to determine the effects of test taking on a person's vocational aspirations, self-understanding, and other criteria. In response, developers and publishers have made multiple revisions to cope with criticism and new data, although no consensus was ever reached on a definition of sex bias or equity in interest measurement.

Hard times have also had an impact on test construction by stimulating the creation of inventories with low development costs. If inventories with high development and maintenance costs such as the SCII or KOIS are no more helpful than inventories with low development and maintenance costs such as the SDS or its clones, then inventories with brief homogeneous scales become attractive. This trend may have been reinforced by scattered and obscure investigations which show that short scales can be developed whose construct validity and retest reliability approximate those obtained by complex and lengthy scales (Meehl, 1972).

Several other influences appear to have accelerated the shift to simpler inventories and a concern with their effects. The need for career services has expanded at a rate that professionals could not or would not cope with so that people with little or no appropriate training have rushed in. Consequently, they found simpler inventories easier to use. In addition, graduate students in counseling psychology, the largest group of professional users of interest inventories, have over the last several decades become more interested in therapy than in vocational tests. Consequently, a test's apparent influence is seen as more important than its psychometric properties.

This increase in test users, whose training ranges from no special training to the Ph.D. in psychology or education, has spawned a large unsophisticated market for vocational tests. And among this growing army of self-taught career counselors is a burgeoning group of entrepreneurs who use bits and pieces of popular inventories to create brief inventories of interests, skills, or workbooks. These people are filling the need for career services, because standard inventories and tests are usually not available to them or because they wanted a special work kit of their own.

The shift to simpler inventories and a concern with effects have also been stimulated by the positive effect of immediate feedback when a person takes a self-scored inventory. Immediacy appears to generate satisfaction and acceptance of the results and to relieve practitioners of some scheduling and coordinating problems such as collating client appointments and test results. Simple inventories also cost less to develop and purchase than inventories with occupational keying and complex scoring.

Finally, the movement toward simpler devices has been accelerated by career theories and classifications (Holland, 1973; Roe, 1956) that use a small number of scales and categories to assess interests and to organize occupational data. The Holland typology led to the SDS and the reorganization of the SVIB as the SCII (Campbell & Hansen, 1981). The Roe theory led to Lunneborg's (1975) Vocational Interest Inventory. Before this theoretical and classificatory work, developers were faced with ad hoc classifications or infinite scale building. Now we have several models for establishing validity generalizations for a limited number of occupational categories and the kinds of interests associated with each. Got-

tfredson, Holland and Ogawa (1982) have applied the Holland typology to the *Dictionary of Occupational Titles* (U.S. Department of Labor, 1977) so that the SDS and similar inventories can be coordinated with 12,099 occupations. McCormick (1979) has also developed an interest inventory and a parallel occupational classification. These classification systems provide practical tools for interpreting inventories, for applying them to nearly all occupations, and for the study of person-job interactions.

The current status of interest testing is also characterized by an increasing use of inventories in computerized career exploration and assessment systems and in micro-computers employed by individuals and institutions. This usage is expected to continue to increase rapidly and to become more decentralized. At present, most usage entails scoring and the production of interpretative reports. One of the most helpful aspects of these developments is that some systems can cumulate test records and can revise scales at a rapid rate. Another positive outcome is the high ratings that computers receive relative to professional counselors (Wagman, 1980; Wagman & Kerber, 1980). The outcomes of these developments appear unusually unpredictable. At this time, computerized test administration and interpretation competes with as well as supplements the work of counselors, and these new services also compete with and supplement printed materials produced by publishers.

Last, the current collection of interest inventories still follows one of two measurement strategies: The first strategy entails an empirical comparison of a person's responses to a reference group. The SCII and KOIS accomplish this task with different methods, but both involve a comparison of a person's responses with a number of occupational samples. These well-established models have also been used in Clark's Minnesota Vocational Interest Inventory (1961) and Johansson's (1982) Career Assessment Inventory. The second strategy entails the development of homogeneous scales defined by factor analyses, typologies, or classifications of interests and occupations. The Holland and Roe typologies have been the most common models for homogeneous scaling. The Self Directed Search (Holland, 1979) and the Vocational Interest Inventory (Lunneborg, 1975) exemplify these applications.

There is a massive amount of data about the merits of individual inventories contained in manuals and in the literature, but there are only a few unequivocal studies of the relative merits of building an inventory following a particular strategy. In addition, a few inventories follow both strategies. For example, the SCII and the Jackson Vocational Interest Survey (JVIS) (Jackson, 1977) use both strategies.

Even a casual review of the evidence implies that different construction strategies result in inventories with similar concurrent and predictive validity. Earlier work (Clark, 1961; Reilly & Echternacht, 1979) suggests that empirical and homogeneous scales have similar predictive and concurrent validity, although

each is occasionally superior to the other for special purposes. Differences are typically very small.

When the evidence for the relative concurrent and predictive validity of inventories constructed by different strategies is added to the evidence for the effects of different inventories on the test-taker, it is difficult to see that one construction strategy is generally superior to another. It is possible that there are some consistent differences for selected purposes, but the kind of research needed to identify such differences has rarely been performed. More important, it is very clear that much higher predictive validity can be obtained by using inventories *and* vocational aspirations in combination. Likewise, the popular strategy of making interest inventories more accessible and interpretable appears equally desirable, and it is a task that has been pursued and studied for only a few years. So I expect the shift in research effort from predictive validity to inventory usefulness to be a trend that will persist for at least a few more years.

## CRITICAL ISSUES AND OPPORTUNITIES

At this time the development and use of interest inventories is faced with four major issues. They include how to make inventories available to more clients versus how to maintain professional incomes and standards, how to create inventories with more valid, influential and satisfying effects, how to insure equity in testing and how to integrate interest testing with other interventions. I found it difficult to label these topics as problems, issues, or golden opportunities. How you perceive these issues depends to some degree on your role (practitioner, publisher, developer, test taker), on your values, on your interpretation of the interest literature, and on your views of career development theory. Consequently, more data will not always create a consensus.

### Improving Psychometric Characteristics

The development of new inventories or the revision of old ones appears to have reached two plateaus. No matter what the method of construction, the reliabilities of individual scales hover around .80 to .90. And the concurrent and predictive validities of diverse inventories appear very similar. This interpretation is not obvious, because the definition of what constitutes a predictive hit or error shifts from one inventory to another, because comparisons of different inventories using the same sample and the same criteria are rare, and because we must piece together a collage of concurrent and predictive studies.

The SCII and KOIS (Kuder & Diamond, 1979) present special problems in evaluation. The SCII uses two kinds of scales—homogeneous and occupational. The six homogeneous scales (GOT) are analogous to the six scales of the SDS



and the UNIACT (Lamb & Prediger, 1981). A review of the manuals for these inventories reveals very similar hit rates for similar and dissimilar criteria. The occupational scales of the SCII have predictive hit rates (Hansen & Swanson, 1983) that closely resemble those for inventories that use homogeneous scales.

Studies of the concurrent and predictive validity of the KOIS also yield hit rates that resemble those of other inventories (Zytowski, 1976). Occasionally, an inventory will stand above the pack, but a careful review will usually disclose a propitious sampling: predictions were performed for very divergent occupations, a rectangular distribution of subjects was employed, small samples in selected categories were omitted, or only a single kind of interest was studied (O'Neil, Magoon, & Tracey, 1978).

It is conceivable that there are some significant differences in inventory validities, but they are difficult to cull out of the literature. If it is correct to assume that we have reached plateaus of retest reliability and predictive validity, we could instead concentrate on making inventories more accessible and more useful to the test-taker. Our psychometrically-oriented colleagues will prefer to spend more time searching for better scaling techniques. However, they will have difficulty reaching the levels of predictive validity attained by combining inventories and vocational aspirations. For example, interest inventories that use a six-variable structure have concurrent or predictive validity hit rates of 39 to 55%. In contrast, when a person's vocational aspiration and interest profile share the same category (person aspires to teacher, and his or her social or educational score is the high point of the profile) predictions of correct identification of future aspiration or occupation range from 60 to 85% for intervals of 1 to 11 years for the SDS, SVIB, or VPI (Bartling & Hood, 1981; Borgen & Seling, 1978; Holland & Gottfredson, 1975; Touchton & Magoon, 1977; Holland & Lutz, 1968). Other work (Holland, Gottfredson, & Nafziger, 1975) implies that the congruency of aspiration and interests is an alternative measure of identity (Holland, Gottfredson, & Power, 1980) and may be a useful predictor of career stability.

### Increasing Inventory Influence

An alternative to the search for better scaling techniques is to continue to make inventories more influential and useful. Nearly all inventories in recent years have been revised to make them easier to interpret and apply. Most publishers have prepared more extensive interpretive materials and have tried to link inventory results to a wide range of occupational information or interventions. They have also attempted to increase the exploration effect by using balanced scales, male and female norms for the same occupation, more occupational scales, norms and raw scores for the same homogeneous interest scale, or by providing comprehensive occupational classification systems that are related to interest profiles. These revisions and the evaluative studies of the effects of

interest inventories and other interventions point to a technology of career instruction or development that may be more productive than the polishing of inventory scales. In short, evaluative studies should lead to inventories that are more influential and practical, although their predictive validity and retest reliability may remain unchanged.

### Giving Psychology Away

The creation of inventories that are more useful to the client also creates some professional conflict. "Giving psychology away" is consistent with our professional ideals. Everyone should have access to an interest inventory if they so desire. Unfortunately, the advent of hard times and self-scored inventories threatens the livelihood of selected professionals. Brief inventories that have obtained positive evaluations threaten some professionals by destroying the mystic of testing and imply the usefulness of inexpensive psychological services. But the old test standards frown on testing without the services of a professional. This situation for psychology is analogous in several ways to the use of cancer, diabetes, and pregnancy diagnostic tests for home use that bypass the family physician, but do not suggest a special treatment.

Research studies that compare the diagnostic services of physicians and do-it-yourself diagnostic kits support the value of many of these simple tests. Likewise, there are at least three experimental evaluations (Avallone, 1974; Krivatsky & Magoon, 1976; Nolan, 1974) that suggest that the usefulness of an interest inventory equals the value of seeing a professional career counselor (i.e., one with a Ph.D. in counseling psychology). In two other experiments (Wagman, 1980; Wagman & Kerber, 1980), an interactive computer system for personal counseling produced positive effects that lasted one month. And, some clients reported that they felt more at ease (42%) and more independent (45%) on the computer than if they saw a counselor. These evaluations and other evaluations imply that interest inventories could be made available to more people with no more expectation of harm or disservice than they would receive from a professional.

### Insuring Equity

Perhaps the most complex, difficult, social-emotional issue is how to insure equity in interest inventory assessments within the constraints of current beliefs, knowledge, test standards, and financial resources. The 10-year controversy about the sex biases in interest testing illustrates the difficulties in arriving at a consensus.

The controversy began in the period 1971 to 1973, when individuals, groups, committees, and commissions charged that interest inventories served to keep

women and men in traditional occupations (Diamond, 1975). The belief was that interest inventories by virtue of their items, instructions, interpretative materials, scoring and normative procedures helped to maintain the sex-segregated character of the work force. This belief was reinforced by numerous plausible hypotheses and some plausible data. For example, when males and females of any age take interest inventories, the distributions of suggested occupations for males and females are usually divergent (Lamb & Prediger, 1981, p. 29). According to interest inventories a small proportion of females have skilled trades and factory-oriented interests, whereas males have those interests in large proportions.

Now, after more than 10 years of discussion and research, there is still no obvious consensus (Tittle & Zytowski, 1978). Perhaps the only area of agreement is that we should be concerned with the influence of inventories on females and all other groups. There is also much disagreement about the interpretation of the more than 100 empirical studies that have evaluated the outcomes of different normative procedures, scaling techniques, special directions, interpretive materials, or the pre-post experiments on the vocational aspirations of females and males.

For example, Zener and Schnuelle (1976) stimulated more than 25 experimental evaluations of the effects of different inventories and other vocational interventions. These investigations suggested that all inventories had beneficial effects on one or more criteria and that the effects appeared similar for inventories constructed and interpreted in different ways.

However, test-taker goals were related to the outcomes of testing in only one investigation (Power, Holland, Daiger, & Takai 1979). That study was important, for it made clear that high school students' expectations for test-taking varied by sex and were not always in accord with the typical counselor goal of exploration. Both boys and girls wanted most of all "reassurance" about a vocational aspiration they already had. "Wanting more alternatives" ranked well below this desire. High school girls wanted more or fewer options to the same degree. Boys wanted fewer options rather than more. Although this study was only a beginning, it supports the experience of practitioners that their clients come to test-taking with different goals. Future evaluations should link test-taker goals to the outcomes of testing. In the Power et al. experiment the necessary analyses were performed, but the N was too small for dependable results.

Another program of studies by Prediger and his colleagues (Lamb & Prediger, 1981) provides another kind of evidence that should lead to a more comprehensive picture of the test-taking experience. Stimulated by a dissertation (Rayman, 1976), this group has demonstrated in numerous studies that it is possible to design an inventory that will suggest similar distributions of vocational alternatives to females and males. They accomplish this by writing items that females and males respond to at about equal rates. The assumption is that this kind of construction will encourage females and males to explore a greater range of

options. This is a provocative idea that should be tested by a demonstration that inventories constructed by the balanced scale method have more exploratory effect than inventories developed by more conventional methods. So far, no one has demonstrated that any inventory has more exploratory effect than another. To the contrary, the exploratory effects of the SCII (Cooper, 1976), SDS (Holland, 1979), Kuder (cited in Zytowski, 1977) and ACT IV (Prediger, McLure, & Noeth, 1976) appear to be very similar; negligible differences are obtained if any.

There are only a few studies of the effects of editorial or psychometric revisions. Experimental revisions of items to determine the effects of gender-neutral words indicate that such revisions have little or no effect on item responses (Boyd, 1976; Gottfredson, 1976; Holland & Gottfredson, 1976). Extensive liberalized directions have failed to increase a person's satisfaction with the outcomes of taking the SDS and have produced only minor profile differences (Siebel & Walsh, 1977). In addition, extensive liberalized directions have failed to increase the number of options or the number of nontraditional options a college woman considers (Lawler, 1977).

In short, it is difficult to demonstrate that any inventory characteristic has a clear impact on the test-taker. One experiment (Holland, Takai, Gottfredson, & Hanau, 1978), of borderline statistical significance, does imply that if an inventory is arranged so that the test-taker can comprehend the scoring, and is provided an instructional booklet as well as a booklet of many options, then it is rated higher on several criteria by test-takers than an inventory with disguised scoring, few options, and no instructional booklet. Hardly a surprise.

I assume that current research activity has dwindled to an occasional lonely article because researchers have slowly realized that only a few read or care about the controversy. We cannot agree because we do not share a common set of beliefs about the purposes of career counseling. Some believe that the only purpose is exploration—especially the exploration or consideration of non-traditional occupations. Others, including myself, believe that career counseling has at least three purposes—exploration, reassurance, and self-understanding—and that these goals vary from person to person. And, because we do not agree on the goals of career counseling and testing, we do not evaluate the research literature in the same way. There are a host of other disagreements: psychometric or internal definitions of bias versus experimental evaluations of actual effects on the test-taker; different views of sex-role socialization (it is an unmitigated evil to some and is celebrated by others); and finally, different goals for interest inventories—social action interventions versus assessment devices.

This controversy and its ambiguous outcomes could be repeated for any new group—retirees, reentry women, Hispanics, blacks, and so on. Somehow we need a strategy of test evaluation and revision that would be generally fair, practical and scientifically sound.

## DESIRABLE DEVELOPMENTS

Now I outline some desirable or needed developments for improving the quality and usage of interest inventories. Many of these developments are underway; and others are potential strategies for which there are sometimes data. Despite my doubts about finding new methods for increasing the validity or reliability of current inventories, only a few inventories appear to take full advantage of current knowledge. There is much room for improvement with standard scaling and validation research. Likewise, if conventional inventories are to compete with vocational workbooks, card sorts, and assorted homemade interest assessments, developers must continue the attempt to make interest inventories more amenable to a wider range of professional and nonprofessional users.

### Improving Psychometric Characteristics

Although a few inventories appear to have reached plateaus of retest reliability and predictive validity, the majority could probably reach the same plateaus with a more active research effort. Many deficiencies could be easily remedied in the case of inventories with computerized scoring services in which substantial data collection is possible. Representative or not, such data would usually be superior to the very small homogeneous samples reported in many test manuals. In addition, Gottfredson, Holland, and Holland (1978) and others have demonstrated how unrepresentative data can be weighted to approximate general population data.

The usefulness of homogeneous and occupational scales needs more examination with more comprehensive occupational samples: occupations of high and low status, and more heterogeneous samples. So far, most comparisons of these scaling methods have been performed on very homogeneous samples—primarily skilled trades or military technical specialties. One or more comprehensive studies might settle the relative merits of these scaling methods. At this time, the evidence implies small differences, but different criteria favor different methods. In short, it looks like a tie.

A variation of the homogeneous versus occupational keying question is how to distinguish among occupations or occupational specialties. Strong (1943) employed a men-in-general group with considerable success, but had difficulty in distinguishing among skilled trades and lower level occupations. Kuder (1966) used the lambda technique, which allowed him to omit a men-in-general group, and obtained efficient discriminations. And inventories that use homogeneous scales simply compare different occupations on a particular set of scales. Each method has produced impressive results, but it is still not clear that one method is generally superior, because there has been no unequivocal study in which all three methods were applied to the same occupational samples. Here another

evaluation would be helpful, although a review of the literature suggests that the differences are again small.

The next step in inventory refinement may be to explore Norman's (1972) strategy for better psychiatric diagnosis; that is, see interest assessment as a "multi-stage, sequential, and branching . . . enterprise in which the discriminations that one must make at one stage depend upon what . . . has already been learned." He uses some old interest inventory research to demonstrate that the selection of an appropriate reference group makes it possible to differentiate among psychologists in different specialties (Kriedt, 1949) and among non-professional occupations (Clark, 1961). This kind of work was abandoned a long time ago—probably because the sampling problems are difficult, expensive, and long-term, but Norman's lucid discussion outlines a persuasive strategy for new work.

Most inventories could be improved by a more systematic representation of the labor force. For example, some inventories sample largely professional level occupations; whereas others sample only the skilled trades. Even if a developer can justify the use of a single subsample, the sampling of the most populous occupations according to census data has merit. Occupational keys for every occupation are not feasible even with unlimited funds. In short, an inventory should represent the occupational domain covered in a more explicit and rational way. The same prescription applies to inventories with homogeneous scales. In this instance, the item content should be clearly related to the particular occupational classification scheme employed.

More explicit research might be performed to establish the "validity generalizations" implied by the General Occupational Theme (G.O.T.) scales of the SCII and the scales of the SDS. For example, if a scale is validated for a few occupations in an occupational category, the scale is probably valid for other occupations in the same category. The required research would entail validity studies for unexamined occupations in the same category to validate both the scales and the occupational classification. As data accumulate, scales and classifications can be revised so that the need for new validity studies will decrease. Some of the needed research has already been performed (Gottfredson et al., 1982), but a more explicit effort would accelerate progress for all inventories that use classifications to interpret inventory scales and profiles.

Developers might also consider the incorporation of aspirational and work history data within the inventory itself or in the interpretive materials. The inclusion of these materials will increase predictive validity substantially and should increase self-understanding. It appears more advantageous to exploit the virtues of self-expressed intentions and work history in combination with inventory scores than it would be to ignore the value of their joint use.

Finally, everyone might profit from reading David Campbell's (1972) book chapter, "The Practical Problems in Revising an Established Psychological

Test.” His discussion of user acceptance (users may slow revisions), technical ignorance (we never have enough knowledge), and administrative arrangements (royalties, funds, and responsibilities of authors and publishers) is a helpful and amusing account of the problems that all developers face, but which are rarely discussed in the measurement literature.

### Increasing Inventory Influence

Perhaps the most promising development would be to accelerate the attempts to make all inventories more useful to the test-taker. The revision of inventories and interpretive aids that began in the 1970s and the investigations of the effects of these revisions are moving us toward a psychology of instruction in career assistance. The simplification of interpretive profiles and the widespread use of homogeneous scales has increased communication between counselor, test-taker, and inventory, and it has probably increased the reliability of test interpretation.

Despite some work on the before-and-after effects of whole inventories (Holland, 1979), there has been relatively little investigation of the actual impact of different interpretive materials or formats. Likewise, most psychometric revisions—separate sex and combined sex norms, scales formed from balanced items (preferred about equally by females and males)—have been examined for concurrent or predictive validity but rarely for their effect on the client. Generally, we do not know if the typical client comprehends and uses the interpretive materials, notices the difference between various scaling methods, or even cares.

If we are to create more useful inventories, more programmatic and comprehensive research is required. The typical research about a test’s psychometric characteristics is necessary but not sufficient. That work needs to be supplemented by more evidence about the effects of different interpretive materials, different scaling procedures, the effects of test-taking on a person’s vocational aspirations, and so on. The assumption that different scaling techniques, formats, and interpretive materials lead automatically to different test-taker outcomes is not supported by the data.

For example, when the Kuder, SCII, and the SDS were administered in different orders to every student in three sections of a career course taught by three different instructors and rated by students on seven criteria (having more vocational alternatives, being reassured, having more self-understanding, and so on), there were no significant differences among the student ratings of these dissimilar inventories (Rayman, 1983). These and other experimental results (O’Neil, Price, & Tracey, 1978) make clear that psychometric and rational analyses of an inventory’s impact are grossly inaccurate signs of actual impact.

It is unlikely that any journal would publish work of this kind or that publishers would share unpublished research. On the other hand, developers and publishers could be encouraged to substantiate claims of usefulness or test-taker



effects in manuals or technical reports. It is conceivable that this kind of research might be published in one or more educational journals because instruction about one's interests is clearly a facet of education as well as career assistance.

The development of self-scored inventories has made inventories available to more clients and more diverse populations. This trend could be accelerated by the reduction of test-taker scoring errors and more assessments of the consequences of self-testing. So far, self-scored inventories compare favorably with professional counselors in three investigations (Avallone, 1974; Krivatsy & Magoon, 1976; Nolan, 1974). In addition, self-scored inventories are relatively neutral or unobtrusive assessment devices. Pallas et al. (1983) found that the SDS was rated as less distressing than 5 of 6 common experiences: "filling out income tax forms, visiting my doctor, taking the written drivers test, or worrying about making ends meet." The SDS was rated as equal to "trying to locate something in the yellow pages." It is reasonable to expect similar ratings for most interest inventories, because they have similar effects on the test-taker.

Self-scored forms of most inventories are possible and may be necessary for populations that require more rapid or inexpensive service. Because self-scored inventories and the research associated with their influence has removed much of the mystification in interest assessment, the old test standards (American Psychological Association, 1974) are out of phase with this recent development, and the recent drafts of the proposed test standards (American Psychological Association, 1984) take a similar obstructive stance by requiring professional supervision.

### Insuring Equity

Insuring equity in interest assessment is a goal that most people endorse, but we lack a consensus on how to achieve that goal for the reasons I outlined earlier—differences in beliefs, values, purposes of interest testing. Accordingly, I have regarded the controversy about the alleged sex biases of interest inventories as not resolvable—at least by the collection of more data. For the same reasons, the potential controversies about the biases of inventories for blacks, Hispanics, aged and other groups have also appeared irresolvable.

Now I can imagine several strategies for coping with these controversies that might lead to more equitable inventory design and usage and that might attract a consensus. The main strategy would be to continue the work in which we are acquiring a more comprehensive and explicit account of what happens when a person takes an inventory. In evaluative studies, inventory purposes, test-taker and counselor goals could be linked to a full range of an inventory's psychometric characteristics and interpretive materials as well as counselor behaviors. Such investigations would help us sort out real from imagined effects; document the effects of revisions planned to increase occupational exploration or self-understanding; document what test-takers want, what they learn, and in follow-up studies, what they remember or how their lives may have been affected.



Although the earlier psychometric and influence studies are a good beginning, a more comprehensive and programmatic set of studies should have more positive outcomes:

1. We would have a short list of inventory characteristics that lead to a particular test-taker outcome.
2. We would have a long list of inventory characteristics that have no effect on the average person.
3. Counselors and test-takers could select inventories that coincide with their goals and values.
4. This information would contribute not only to the development of more helpful interest inventories but also to other interventions such as career courses, career counseling, vocational card sorts, and workbooks.

I see this venture as a win-win-win situation for test-takers, developers, and counselors. If it can be demonstrated that specific inventory characteristics create specific outcomes, then the debates about inventories will be clarified, and counselors can deliver treatments that are more consonant with client goals. If current inventories do not generate differential effects, we can turn to more promising areas of research and development. Finally, no matter what outcomes are obtained—clear or mushy—we will have generated a more rational and comprehensive knowledge of inventory influence that can be readily applied to new questions of equity and old questions of career assistance.

I have emphasized influence studies, because I believe tested effects (external criteria) are more defensible than the growing list of psychometric signs of bias. These special internal analyses form a heterogeneous collection of intermediate criteria that have a special set of ambiguities—especially if a developer performs two or more tests for bias. I am reminded of a news item in which two MIT aeronautical engineering students demonstrated in a computer analysis that the common bee should not be capable of flight.

Having said this, I would still advocate a review of test items for extreme endorsement rates according to age, sex, SES, race and other attempts (judgmental and statistical) to identify items that did not function in the same way across different groups. My experience is that tests of this kind reidentified bad items identified earlier by traditional item analysis.

The problems inherent in the use of internal analyses as a means for detecting bias are apparent in a recent *Handbook of Methods for Detecting Test Bias* (Berk, 1982). Sixteen authors outline more than sixteen methods (sometimes contradictory) of detecting bias. This book is a valuable source of defensive maneuvers for developers and publishers. At the same time, Burrill's (1982) chapter, "Comparative Studies of Item Bias Methods" summarizes the ambiguities and practical difficulties that need to be overcome before some useful consensus is obtained.

## PROBABLE DEVELOPMENTS

My sketch of desirable developments is only one of several. Different people would have a different list of desirable developments or different emphases for the same topics. On the other hand, a list of probable developments based on the strength and popularity of particular current developments should result in more agreement and is probably more predictive of what will happen.

Two current developments appear likely to survive and flower: (a) the creation of inventories that are easier to develop, interpret, and use with a wider range of people, and (b) the pursuit of evaluative and experimental research that may lead to a clearer understanding of how inventories achieve their effects. These trends have little opposition and are fueled by the demands of special groups, the needs of publishers, developers, practitioners and researchers.

In contrast, future work on equity in testing looks like a no-win situation for developers, publishers, and reformers who have already been through 10 years of controversy in which both defenders and reformers have come away bruised and confused. Similar controversies about the interests of reentry women, retirees, blacks, and Hispanics do not appear more rewarding, for they would contain all the ambiguity and complexity that the sex bias controversy exhibited. Instead I believe the question of equity should shift from vague or debatable beliefs about equity to the study of particular effects so that people can then select inventories according to their demonstrated influences. Likewise, the development of inventories with more construct validity and reliability appears unlikely, because inventories developed by long-term expensive methods are not clearly superior to inventories developed from vocational theory or homogeneous scaling techniques.

This interpretation is strengthened by Meehl's (1972) long-term attempts to build a better MMPI. His work suggests that ideal scales are brief (15–20 items), homogeneous, and easily interpreted by both clients and theorists. In addition, ideal scales should function effectively "in subpopulations homogeneous with regard to age, sex, education, IQ, race, social class. . . ." Much of the research about old and new interest inventories leads to the same interpretations.

### More Practical Inventories

The development of more practical inventories has been underway for more than a decade and appears to be accelerating. By "practical" I mean inventories that are inexpensive to develop, interpret and use, but are also scientifically sound. Such inventories are characterized by homogeneous scales, self-scoring, many occupational options, and readable, comprehensive interpretive materials.

I would also include several other trends in the movement toward more practical inventories. They include the growth of computer-interpreted inventories and the inclusion of standard inventories within computer-assisted career

systems. The multiple attempts to make inventories more helpful to females, people with skilled trades interests, and to older Americans are also attempts to make inventories more useful to old and new clients. The provision of more vocational information and treatment ideas in interpretive materials as well as the publication of inventories in comprehensive workbooks or along with a compatible set of related ability, values or other tests is another facet of this trend. We need research evaluations that compare the effects of workbooks with more labor-intensive treatments such as one-to-one career counseling, seminars, and courses.

### Inventories as Interventions

The study of inventories as interventions should continue to increase because it is consistent with the attempts to create more useful inventories as well as with the diffuse research and theoretical effort devoted to increasing the quality of career assistance. This kind of work will not proceed at a rapid rate, for it has the support or interest of only a few publishers, developers, and researchers. It is easier to perform psychometric analyses and revise interpretive materials without testing their impact, for the research required to assess inventory effects or special revisions is characterized already by some negative results, difficulties of execution, and a lack of publication outlets. On the positive side, this research requires little funding, but like many evaluations, it requires considerable cooperation, coordination, and negotiation. Consequently, one research team is lucky if they can do one evaluation a year. In the meantime, critics have developed a new list of outcome criteria.

## INNOVATIONS AND CONSTRAINTS

Finally, it is important to acknowledge that inventories will be affected by future technological and theoretical events as well as by multiple environmental constraints that will alter any predictions of desirable or probable developments. Some recent events illustrate these unpredictable forces.

A recent Ph.D. thesis (Monahan, 1983) has demonstrated that a simple 60-item inventory of a student's environmental preference was more predictive of a student's current vocational aspiration (81% agreement) than the Vocational Preference Inventory (Gottfredson, Holland, & Holland, 1978), (55%) composed of 84 items. This work implies a new assessment strategy that is probably more efficient, and perhaps more conducive to learning. It could shift all controversy about bias to questions of objective environmental descriptions because test-takers express their preference for a variety of occupational environments rather than their preferences for different activities, occupational titles, and other traditional test items. In another innovation, Prediger (1982) has reduced the

measurement of interests to two dimensions and has shown how all occupations can be displayed in a single plane divided into 12 regions.

Both innovations represent important developments, but both face strong environmental constraints—largely the traditional beliefs that professionals have about interest measurement, and the aversion that many people have to change. For example, the major revisions of the SCII have left some old timers unhappy. The publication of the SDS in 1971 with self-scoring and transparent scales also offended some professionals.

The proliferation of computers and microcomputers has the potential for making everyone a test publisher. This task can be accomplished by simply inserting an inventory into a computer and by adding a scoring and interpretive program, so that the computer becomes the publisher. These new publishers cannot be easily monitored by professional associations, traditional test publishers, or the public. When these technological developments are coupled with the proliferation of self-help work books, card sorts, and plausible but home-made interest inventories, they create a threat to traditional test publishers who are coerced to some degree by the old test standards.

The new test standards are still in revision, but the first four drafts would have required test publishers to conduct more research prior to publication than many, perhaps most, publishers would risk. The proposed standards imply not only more financial but also more legal risk than in the past. Old inventories should find compliance easier, but the legal risk may be similar for old and new inventories. We shall have to wait and see, but in no way can the proposed standards be regarded as a stimulus to innovation. Some compromises are needed to thread new and old inventories between idealistic standards and no standards. These might include selected exemptions for new inventories for a limited time and similar exemptions for inventories with small sales but again for a limited time. Likewise, exemptions should be possible for inventories developed for special populations, or it should be possible to advise professionals not to use inventories with selected groups to avoid the need to turn inventories into omnibus devices.

## CONCLUSION

I have attempted to summarize the current status of interest testing, to outline some critical issues and to project what may or should happen next. Because some of my interpretations of this literature deal with very controversial topics, it is desirable to emphasize some interpretations, hopes, or trends for which there is less controversy.

First, I conclude that the need to understand how inventories create their effects is still with us and is an ideal research enterprise for developing more helpful inventories, for coping with equity questions, and for understanding inventories as interventions.

Second, I conclude that a greater effort needs to be made to stimulate authors and publishers to reach higher levels of validity where there has been little or insufficient research. Although I am pessimistic about attempts to increase predictive validity by new psychometric methods, many developers and publishers have not taken advantage of old methods. I have doubts about more psychometric research to develop scales with more validity, but I still hope that a limited effort would be devoted to this strategy. My rationale for these interpretations is two-fold: (a) More than 40 years of scaling research has not resulted in higher levels of predictive validity, but (b) only 10 years of research have resulted in inventories that are more interpretable, easier to score, and have lead to a clearer understanding of interest inventories as career treatments. There comes a time in research when you must cut your losses and invest in more promising strategies or at least redistribute your resources. That time seems now.

### ACKNOWLEDGMENT

I am indebted to Fred H. Borgen, Gary D. Gottfredson, Linda S. Gottfredson, Jo-Ida Hansen, Thomas M. Magoon, and Jack R. Rayman for their counsel about earlier drafts of this chapter.

### REFERENCES

- American Psychological Association. (1974). Standards for educational and psychological tests and manuals. Washington, D.C.: Author.
- American Psychological Association. (1984). Joint technical standards for educational and psychological testing. (Fifth draft). Washington, D.C.: Author.
- Association for Measurement and Evaluation in Guidance. (1973). AMEG commission report on sex bias in interest measurement. *Measurement and Evaluation in Guidance*, 6, 171–177.
- Avallone, V. L. (1974). A comparative study of the effects of two vocational guidance systems: The Self-Directed Search and a traditional vocational guidance model. *Dissertation Abstracts International*, 35, 2670A–2671A. (University Microfilms No. 74–24, 472).
- Bartling, H. C., & Hood, A. B. (1981). An 11-year follow-up of measured interest and vocational choice. *Journal of Counseling Psychology*, 28, 27–35.
- Berk, R. A. (Ed.). (1982). *Handbook of methods for detecting test bias*. Baltimore: Johns Hopkins Press.
- Borgen, F. H., & Seling, M. J. (1978). Expressed and inventoried interests revisited: Perspicacity in the person. *Journal of Counseling Psychology*, 25, 536–543.
- Boyd, V. S. (1976). Neutralizing sexist titles in Holland's Self-Directed Search: What difference does it make? *Journal of Vocational Behavior*, 9, 191–199.
- Buros, O. K. (Ed.). (1978). *The Eighth Mental Measurements Yearbook* (pp. 1535–1649). Highland Park, NJ: Gryphon Press.
- Burrill, L. E. (1982). Comparative studies of item bias methods. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias* (pp. 161–179). Baltimore: Johns Hopkins Press.
- Campbell, D. P. (1972). The practical problems of revising an established psychological test. In J. N. Butcher (Ed.), *Objective personality assessments* (pp. 117–130). New York: Academic Press.

- Campbell, D. P. & Hansen, J. C. (1981). *Manual for the SVIB-SCII*, 3rd Edition. Stanford: Stanford University Press.
- Clark, K. E. (1961). *The vocational interests of nonprofessional men*. Minneapolis: University of Minnesota Press.
- Cooper, J. F. (1976). Comparative impact of the SCII and the vocational card sort on career salience and career exploration of women. *Journal of Counseling Psychology*, 23, 348–352.
- Diamond, E. E. (Ed.). (1975). *Issues of sex bias and sex fairness in career interest measurement*. Washington, D.C.: National Institute of Education.
- Dolliver, R. H. (1969). Strong Vocational Interest Blank versus expressed occupational interests: A review. *Psychological Bulletin*, 72, 94–107.
- Evans, J. R., & Rector, A. P. (1978). Evaluation of a college course in career decision-making. *Journal of College Student Personnel*, 19, 163–168.
- Gottfredson, G. D. (1976). A note on sexist wording in interest measurement. *Measurement and Evaluation in Guidance*, 8, 221–223.
- Gottfredson, G. D., & Holland, J. L. (1978). Toward beneficial resolution of the interest inventory controversy. In C. K. Tittle & D. G. Zytowski (Eds.), *Sex-fair interest measurement* (pp. 43–51). Washington, D.C.: National Institute of Education.
- Gottfredson, G. D., Holland, J. L., & Holland, J. E. (1978). The Seventh revision of the Vocational Preference Inventory. *Psychological Documents*, 8, 98 (Ms. 1783)
- Gottfredson, G. D., Holland, J. L., & Ogawa, D. K. (1982). *Dictionary of Holland Occupational Codes*. Palo Alto, CA: Consulting Psychologists Press.
- Hansen, J. C. & Swanson, J. L. (1983). Stability of interests and the predictive and concurrent validity of the 1981 Strong-Campbell Interest Inventory for college majors. *Journal of Counseling Psychology*, 30, 194–201.
- Holland, J. L. (1973). *Making vocational choices*. Englewood Cliffs, NJ: Prentice-Hall.
- Holland, J. L. (1979). *Professional manual for the Self-Directed Search*. Palo Alto: CA: Consulting Psychologists Press.
- Holland, J. L., & Gottfredson, G. D. (1975). Predictive value and psychological meaning of vocational aspirations. *Journal of Vocational Behavior*, 6, 349–363.
- Holland, J. L., & Gottfredson, G. D. (1976). Sex differences, item revisions, and the Self-Directed Search. *Measurement and Evaluation in Guidance*, 8, 224–228.
- Holland, J. L., Gottfredson, G. D., & Nafziger, D. H. (1975). Testing the validity of some theoretical signs of vocational decision-making ability. *Journal of Counseling Psychology*, 22, 411–422.
- Holland, J. L., Gottfredson, D. C., & Power, P. G. (1980). Some diagnostic scales for research in decision-making and personality: Identity, information, and barriers. *Journal of Personality and Social Psychology*, 39, 1191–1200.
- Holland, J. L., & Lutz, S. W. (1968). The predictive value of a student's choice of vocation. *Personnel and Guidance Journal*, 46, 428–436.
- Holland, J. L., Takai, R., Gottfredson, G. D., & Hanau, C. A. (1978). Multivariate analysis of the effects of the Self-Directed Search on high school girls. *Journal of Counseling Psychology*, 25, 384–389.
- Jackson, D. N. (1977). *Manual for Jackson Vocational Interest Survey*. Port Huron, MI: Research Psychologists Press.
- Johansson, C. B. (1982). *Manual for The Career Assessment Inventory*. Minneapolis: National Computer Systems.
- Kriedt, P. H. (1949). Vocational interests of psychologists. *Journal of Applied Psychology*, 33, 482–488.
- Krivatsy, S. E., & Magoon, T. M. (1976). Differential effects of three vocational counseling treatments. *Journal of Counseling Psychology*, 23, 112–118.
- Kuder, G. F. (1954). Expected developments in interest and personality inventories. *Educational and Psychological Measurement*, 14, 265–271.

- Kuder, G. F. (1966). The Occupational Interest Survey. *Personnel and Guidance Journal*, 45, 72–77.
- Kuder, F., & Diamond, E. E. (1979). General Manual for the Occupational Interest Survey. Chicago, IL: Science Research Associates.
- Lamb, R. R., & Prediger, D. J. (1981). *Technical report for the unisex edition of the ACT Interest Inventory*. Iowa City, IA: American College Testing Program.
- Lawler, A. C. (1977). *Career exploration with women using the Non-sexist Vocational Card Sort and the Self-Directed Search*. Unpublished doctoral dissertation, College of Education, University of North Carolina, Chapel Hill.
- Lunneborg, P. W. (1975). *Manual for the Vocational Interest Inventory*. Seattle: University of Washington, Educational Assessment Center.
- McCormick, E. J. (1979). *Job analysis: Methods and applications*. New York: AMACOM, American Management Association.
- Meehl, P. E. (1972). Reactions, reflections, projections. In J. N. Butcher (Ed.), *Objective personality assessment* (pp. 131–189). New York: Academic Press.
- Monahan, C. J. (1983). *A decision theoretic approach to the predictions of vocational preference*. Unpublished doctoral dissertation, Iowa State University, Ames.
- Nolan, J. J. (1974). The effectiveness of the Self-Directed Search compared with group counseling in promoting information-seeking behavior and realism of vocational choice. *Dissertation Abstracts International*, 35, 195A. (University Microfilms No. 74–16, 569)
- Norman, W. T. (1972). Psychometric considerations for a revision of the MMPI. In J. N. Butcher (Ed.), *Objective personality assessment* (pp. 59–83). New York: Academic Press.
- O'Neil, J. M., Magoon, T. M., & Tracey, T. J. (1978). Status of Holland's investigative personality types and their consistency levels seven years later. *Journal of Counseling Psychology*, 25, 530–535.
- O'Neil, J. M., Price, G. E., & Tracey, T. (1978). The stimulus value treatment effects, and sex differences when completing two interest inventories. Counseling Center Research Report No. 78–2, *Resources in Education*. ERIC Document Reproduction Service, ED 158 208.
- Pallas, A. M., Dahmann, J. S., Gucer, P. W., & Holland, J. L. (1983). Test-taker evaluations of the Self-Directed Search and other psychological tests. *Psychological Documents*, 13, 11. (Ms. No. 2550)
- Power, P. G., Holland, J. L., Daiger, D. C., & Takai, R. T. (1979). The relation of student characteristics to the influence of the Self-Directed Search. *Measurement and Evaluation in Guidance*, 12, 98–107.
- Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, 21, 259–287.
- Prediger, D. J., McLure, G. T., & Noeth, R. J. (1976). Promoting the exploration of personally relevant career options in science and technology (NSF Grant No. SM175–18149 A01). Iowa City, IA: American College Testing Program.
- Rayman, J. (1976). Sex and the single interest inventory: The empirical validation of sex-balanced interest inventory items. *Journal of Counseling Psychology*, 23, 239–246.
- Rayman, J. R. (1983). *The effects of three career courses and selected career treatments*. Unpublished manuscript, Iowa State University, Ames.
- Rayman, J. R., Bernard, C. B., Holland, J. L., & Barnett, D. C. (1983). The effects of a career course and some other popular treatments. *Journal of Vocational Behavior*, 23, 346–355.
- Reilly, R. R. & Echternacht, G. J. (1979). Validation and comparison of homogeneous and occupational interest scales. *Applied Psychological Measurement*, 3, 177–185.
- Roe, A. (1956). *The psychology of occupations*. New York: Wiley.
- Siebel, C. E., & Walsh, W. B. (1977). A modification of the instructions to Holland's Self-Directed Search. *Journal of Vocational Behavior*, 11, 282–290.
- Strong, E. K., Jr. (1943). *Vocational interests of men and women*. Stanford: Stanford University Press.

- Sundberg, N. D. (1977). *Assessment of persons*. Englewood Cliffs, N.J.: Prentice-Hall.
- Tittle, C. K., & Zytowski, D. G. (Eds.). (1978). *Sex-fair interest measurement: Research and implications*. Washington, D.C.: National Institute of Education.
- Touchton, J. B., & Magoon, T. M. (1977). Occupational daydreams as predictors of vocational plans for college women. *Journal of Vocational Behavior*, 10, 156–166.
- U.S. Department of Labor. (1977). *Dictionary of occupational titles* (Fourth edition). Washington, D.C.: U.S. Government Printing Office.
- Wagman, M. (1980). PLATO DCS: An interactive computer system for personal planning. *Journal of Counseling Psychology*, 27, 16–30.
- Wagman, M., & Kerber, K. W. (1980). PLATO DCS, an interactive computer system for personal counseling: Further development and evaluation. *Journal of Counseling Psychology*, 27, 31–39.
- Zener, T. B., & Schnuelle, L. (1976). Effects of the Self-Directed Search on high school students. *Journal of Counseling Psychology*, 23, 353–359.
- Zytowski, D. G. (1976). Predictive validity of the Kuder Occupational Interest Survey: A 12- to 19-year follow-up. *Journal of Counseling Psychology*, 3, 221–233.
- Zytowski, D. G. (1977). The effects of being interest-inventoried. *Journal of Vocational Behavior*, 11, 153–157.
- Zytowski, D. G., & Borgen, F. H. (1983). Assessment (pp. 5–40). In S. H. Osipow & W. B. Walsh (Eds.), *Handbook of Vocational Psychology* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates.



